DOCUMENT RESUME

ED 327 566 TM 015 996

AUTHOR Facione, Peter A.
TITLE The California Critical Thinking Skills Test--College Level. Technical Report #4. Interpreting the CCTST, Group Norms, and Sub-Scores.
PUB DATE 90
NOTE 19p.; For reports 1-3, see TM 015 818-819 and ED 326 584.
AVAILABLE FROM California Academic Press, 217 La Cruz Ave., Millbrae, CA 94030.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Cognitive Tests; *College Students; *Critical Thinking; Deduction; General Education; Higher Education; Induction; Inferences; *Multiple Choice Tests; *Norms; Pretests Posttests; *Scores; Standardized Tests; *Test Interpretation; Thinking Skills
IDENTIFIERS *California Critical Thinking Skills Test (College); Percentile Ranks; Subtests

ABSTRACT

Group norms are provided for the California Critical Thinking Skills Test (CCTST)--College Level, a standardized 34-item multiple-choice test designed to assess the core critical thinking skills associated with baccalaureate general education. The CCTST offers three subtests conceptualized in terms of a national Delphi study on critical thinking. These three subtests--analysis, evaluation, and inference--correlate strongly with each other and the overall CCTST when used as either a pretest or posttest. Subtests are also offered based on the more traditional division of reasoning into "deductive reasoning" and "inductive reasoning." These latter two subtests also correlate strongly with each other and the overall CCTST when used as either a pretest or posttest. Statistical analyses, correlations, and recommended percentile rankings for raw scores are presented in nine tables. These norms were developed on the basis of analyses of 1,673 test forms for representative samples of college students at a comprehensive urban state university during the 1989-90 school year. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED327566

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

PETER A. FACIONE

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

The California Critical Thinking Skills Test -- College Level

Technical Report #4

Interpreting the CCTST, Group Norms, and Sub-Scores

Peter A. Facione
Santa Clara University

c. 1990
CALIFORNIA ACADEMIC PRESS
217 LaCruz Ave., Millbrae , CA 94030

TM015996

The California Critical Thinking Skills Test: College Level

Technical Report #4 --

Interpreting the CCTST, Group Norms and Sub-Scores

by

Peter A. Facione

Santa Clara University

Abstract

Technical Report #4, in a series of four, provides group norms for the **California Critical Thinking Skills Test: College Level**, a standardized testing instrument designed to assess the core critical thinking skills associated with baccalaureate general education. The CCTST offers three sub-tests conceptualized in terms of the recently completed national Delphi study, Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction. These three sub-tests, "Analysis," "Evaluation," and "Inference," correlate strongly with each other and with the overall CCTST. The CCTST also offers sub-tests based on the more traditional division of the reasoning arts into "Deductive Reasoning" and "Inductive Reasoning." Complete statistical analyses, correlations and recommended percentile rankings for raw scores on each of the five sub-tests as well as for the CCTST overall, used either in a pretest or posttest context, are presented in tabularized form in this technical report. These norms have been developed on the basis of analyses of 1673 test forms completed by representative samples of college students during the 1989/90 academic year at a comprehensive urban state university. Technical Report #1 in this series reports on the content validity of the CCTST and its experimental validation during 1989/90. Technical Report #2 describes the concurrent validity of the CCTST in terms of its correlations with SAT-verbal, SAT-math, college GPA, and Nelson-Denny Reading Test scores. Technical Report #3 reports on the relationship between CCTST and four student-related variables: gender, ethnicity, academic major and CT self-esteem.

The California Critical Thinking Skills Test: College Level
Technical Report #4 --

Interpreting the CCTST, Group Norms and Sub-Scores

by

Peter A. Facione
Santa Clara University

Recap of Previous Findings

This Technical Report, the fourth and final in this series, provides detailed statistical information on the five CCTST sub-tests. Three sub-tests are conceptualized in terms of the recently completed national Delphi study, Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction (Facione, 1990 a). These three sub-tests, "Analysis," "Evaluation," and "Inference," correlate strongly with each other and with the overall CCTST, used as either a pretest or a posttest. The same is true of the two CCTST sub-tests, "Deductive Reasoning" and "Inductive Reasoning," which divide CCTST items along that more traditional matrix. Recommended percentile rankings for raw scores on each of the five and for the CCTST overall -- used either as pretests or posttests -- have been developed. The statistical analyses which form the basis for these recommendations were conducted on the 1673 CCTST test forms completed by representative samples of college students enrolled in campus approved critical thinking

courses and control group courses during the 1989/90 academic year at a comprehensive urban state university.

Technical Report #1 in this series discussed the content validity of the CCTST in terms of the conceptualization of CT expressed in Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction as well as the concept of CT grounding the system-wide CT general studies requirement of the California State University. Also, Technical Report #1 described a series of four experiments which indicated that the CCTST is an effective measure of the improvements in the core CT skills of interpretation, analysis, evaluation, inference and explanation which occur as a result of taking a lower division college level CT course. During 1989/90, data was collected on a variety of variables relating to the 20 instructors and the 1196 college students who participated in these experiments. Those studied were either teaching or enrolled in 45 sections of five different courses offered by three departments, (Facione, 1990 c).

Technical Report #2 described the relationship of CCTST results to a number of student-related and instructor-related variables. Critical thinking skills, as measured on the CCTST, can be predicted by a combination of SAT verbal, SAT math, and GPA data with R-square =.41 If CCTST pretest data are included in the regression model the R-square =.71. A college student's age, units of college work completed, and high school subject matter preparation, and an instructor's teaching experience do not contribute significantly to the regression models which predict CCTST posttest results. CCTST results positively correlated with Nelson-Denny reading scores for vocabulary, comprehension, and total score. Non-native English speakers show virtually no gain from CCTST pretest to posttest and, hence, use of the CCTST for non-native English speaking students is counter-indicated. Of six instructor-related factors which are thought to be

related to effectiveness in teaching CT skills, only years of teaching experience and recent experience teaching CT are related, and these in non-linear ways. No evidence was found to support the hypothesis that CT skill development is a natural outcome of baccalaureate education, either in general, or by reference to the control groups, (Facione, 1990 d).

Technical Report #3 examined the CCTST in terms of the possible impact of student gender, ethnicity, academic major and CT self-esteem on CT skill performance. Analyses of pretest data and control group data show that the CCTST is not gender-biased. Statistically significant gender differences emerge only after students complete their college level CT course. ANCOVA also indicated that the CCTST does not favor or disadvantage any particular ethnic or racial group. However, not all groups appeared to benefit equally from having completed their approved college level CT course. While academic major was not a significant factor on the CCTST pretest, scores on the posttest did vary significantly by major. Student CT self-confidence, which appears unrealistically high, did correlate with relative success on the CCTST. However, when SAT and native language were controlled, CT self-confidence was not a significant factor in explaining pretest or posttest results. The emergence of significant differences by gender, ethnicity and major on the CT posttests indicated an urgent need for research on student learning relative to CT curriculum and CT pedagogy, (Facione, 1990 e).

CCTST Pretest and Posttest Norms

In its final form the CCTST is a standardized 34 item multiple choice assessment tool. Twenty of the questions offer four choices, fourteen offer five. For purposes of CT skill assessment, one answer has been designated the superior choice on each question. All distractors ("wrong" answers) were selected by some subjects in the CCTST validation

studies during 1989/90 as well as in the prior years of individual item pilot testing.

To establish stable pretest and posttest norms the largest possible number of subjects was used. Pretest norms are based on the responses of 781 college students who completed the CCTST as a pretest in Feb. 1990 during week one or two of an approved CT course or who completed the CCTST as either a pretest or posttest in the control group (non-CT) course.[1] Posttest norms are based on the responses of 892 college students who completed the CCTST in Nov. 1989 or May 1990 during week 14 or 15 of a three semester unit college level course approved as meeting a campus general studies CT requirement. Table 1 displays pretest and posttest statistics.

Of the 1673 tests evaluated, the top score achieved was a posttest 31 and the lowest a pretest 2. There is room for group movement both above and below both means as well as beyond the outliers of both the pretest and posttest. The statistics on Table 1 and the histographic representation of the curves produced on the pretest and on the posttest on Table 2 indicate that both curves are sufficiently normal.

Table 1

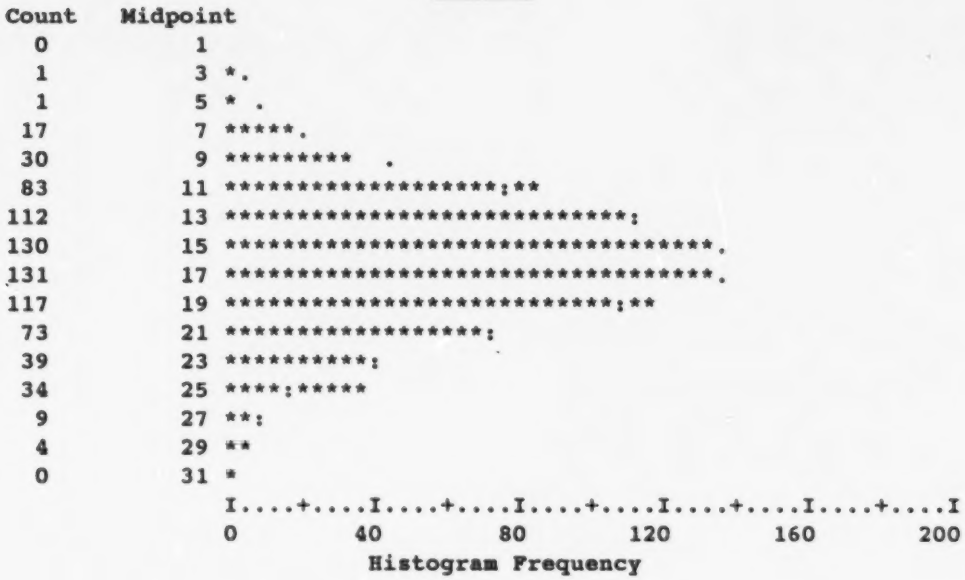
Statistical Analysis of Pretest and Posttest Groupings

PRETEST		POSTTEST	
Mean	15.890	Mean	17.272
Std Err	.159	Std Err	.161
Median	16.000	Median	17.000
Mode	16.000	Mode	15.000
Std Dev	4.457	Std Dev	4.823
Variance	19.862	Variance	23.265
Kurtosis	-.133	Kurtosis	-.368
S E Kurt	.175	S E Kurt	.164
Skewness	.192	Skewness	.136
S E Skew	.087	S E Skew	.082
Range	27.000	Range	28.000
Minimum	2.000	Minimum	3.000
Maximum	29.000	Maximum	31.000
Sum	12410.000	Sum	15407.000
Valid Cases	781	Valid Cases	892

Table 2

Normal Curves for Pretest and Posttest Groupings

PRETEST



POSTTEST

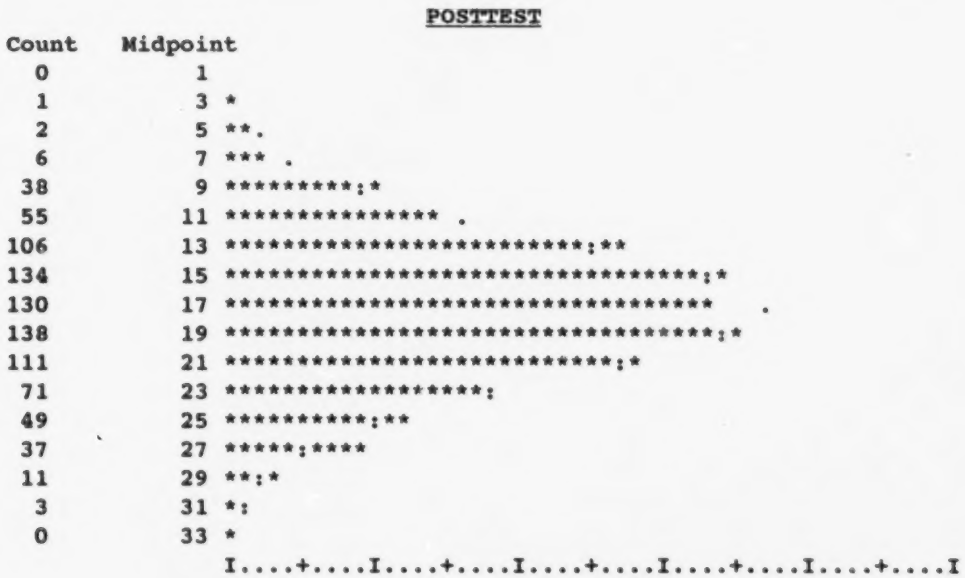


Table 3 represents the percentiles recommended to be associated with each raw score for the pretest and for the posttest. For example, a student who answers 20 correctly

on the pretest would rank in the 86th percentile. The same number correct on the posttest would rank in the 75th percentile. This drop in percentile ranking is to be expected because of the measurable improvement in the group's CT skills.

Table 3

Percentile Rankings for Pretest and Posttest Raw Scores

Number Correct	Pretest Frequency	Cum %	Pretest Percentile	Posttest Frequency	Cum %	Posttest Percentile
1	0	0.0	1	0	0.0	1
2	1	0.1	1	0	0.0	1
3	0	0.1	1	1	0.1	1
4	0	0.1	1	2	0.3	1
5	1	0.3	1	0	0.3	1
6	10	1.5	2	1	0.4	1
7	7	2.4	2	5	1.0	1
8	10	3.7	4	11	2.2	2
9	20	6.3	6	27	5.3	5
10	33	10.5	11	19	7.4	7
11	50	16.9	17	36	11.4	11
12	50	23.3	23	46	16.6	17
13	62	31.2	31	60	23.3	23
14	64	39.4	40	51	29.0	29
15	66	47.9	48	83	38.3	38
16	68	56.6	57	69	46.1	46
17	63	64.7	65	61	52.9	53
18	64	72.9	73	62	59.9	60
19	53	79.6	80	76	68.4	68
20	47	85.7	86	60	75.1	75
21	26	89.0	89	51	80.8	81
22	24	92.1	92	35	84.8	85
23	15	94.0	94	36	88.8	89
24	19	96.4	96	23	91.4	91
25	15	98.3	98	26	94.3	94
26	3	98.7	99	26	97.2	97
27	6	99.5	99	11	98.4	98
28	2	99.7	99	8	99.3	99
29	2	100.0	99	3	99.7	99
30	0	100.0	99	2	99.9	99
31	0	100.0	99	1	100.0	99
32	0	100.0	99	0	100.0	99
33	0	100.0	99	0	100.0	99
34	0	100.0	99	0	100.0	99

To insure a more accurate representation of the population of college students at comprehensive public universities, the norms and percentile recommendations presented in this technical report include the scores of native and non-native English speaking

students. To the extent a given group of persons tested with the CCTST might differ from this norm group on factors predictive of CT skills CCTST users should consider making local modifications in the recommended percentiles. For details consult Technical Reports #2 and #3, (Facione, 1990 d, e).[2]

The Analysis, Evaluation, and Inference Sub-Tests

The items on the CCTST can be divided along either of two theoretical matrices. The first, developed out of the Delphi research, sub-divides the entire CCTST, all 34 items, into three distinct groupings: Analysis, Evaluation and Inference. The second, using a more traditional conceptualization, sub-divides 30 the 34 CCTST items into two groupings: Deductive Reasoning and Inductive Reasoning.

Using the Delphi matrix, items 1-9 fall under the sub-score named "analysis" and relate to the core CT skills of interpretation and analysis. Items 10-13 and 25-34 are grouped under the sub-score of "evaluation" and relate to the core CT skills of evaluation and explanation. Items 14-24 are grouped under "inference" and relate to the core CT skill of inference, (Facione, 1990 c). Thus, on the Delphi matrix each item is included on one, and only one, sub-test; nine items are used for "analysis," fourteen for "evaluation," and eleven for "inference."[3]

Table 4 indicates the correlations between each sub-score and the two others, as well as between each and the overall pretest and posttest scores. Table 5 displays statistical data for each of the three sub-tests and Table 6 indicates the recommended percentile rankings for each in both the pretest and posttest contexts using raw scores. Discretion is recommended in the use of sub-test results and percentile rankings. Use

should be restricted to diagnostic purposes, the evaluation of CT programs, or the assessment of aggregate groups of students, in contrast to summative evaluations made of individual persons.[4]

Table 4[5]

Correlations of Delphi Matrix Sub-Tests

	Analysis		Evaluation		Inference	
	PRE-ANAL	POST-ANAL	PRE-EVAL	POST-EVAL	PRE-INFR	POST-INFR
PRETEST	.6363 n= 601 p=.000	.3941 n= 412 p=.000	.8096 n= 601 p=.000	.5599 n= 412 p=.000	.7634 n= 601 p=.000	.4753 n= 412 p=.000
POSTTEST	.3465 n= 319 p=.000	.6348 n= 872 p=.000	.5101 n= 319 p=.000	.8344 n= 872 p=.000	.5160 n= 319 p=.000	.7867 n= 872 p=.000
PRE-ANAL	1.000	.3283 n= 318 p=.000	.2944 n= 601 p=.000	.2767 n= 318 p=.000	.3177 n= 601 p=.000	.2234 n= 318 p=.000
POST-ANAL	above	1.000	above	.2906 n= 872 p=.000	above	.3627 n= 872 p=.000
PRE-EVAL	above	.2878 n= 318 p=.000	1.000	.5072 n= 318 p=.000	.4083 n= 601 p=.000	.3283 n= 318 p=.000
POST-EVAL	above	above	above	1.000	above	.4415 n= 872 p=.000
PRE-INFR	above	.2580 n= 318 p=.000	above	.4129 n= 318 p=.000	1.000	.4785 n= 318 p=.000
POST-INFR	above	above	above	above	above	1.000

For example, to find the correlation between a sub-score on "Inference" and the CCTST overall used in a posttest context, read "POSTTEST" across to the two columns on the right. The correlation of pretest inference sub-scores is .5160 and the correlation of the posttest inference sub-scores is .7867. Notice that pretest sub-scores correlate more strongly with the CCTST pretest overall score whereas the posttest sub-scores correlate more strongly with the CCTST overall posttest score. To find out how well any one sub-score correlates with another, find the intersection of the appropriate row and column.

Table 5

Statistical Analyses of Delphi Matrix Sub-Tests

	Analysis		Evaluation		Inference	
	PRE-ANAL	POST-ANAL	PRE-EVAL	POST-EVAL	PRE-INFR	POST-INFR
Mean	4.454	4.766	5.406	6.178	6.141	6.349
Std Err	.064	.052	.100	.090	.083	.069
Median	5.000	5.000	5.000	6.000	6.000	6.000
Mode	5.000	5.000	6.000	6.000	7.000	6.000
Std Dev	1.558	1.536	2.449	2.662	2.028	2.052
Variance	2.428	2.361	5.998	7.088	4.112	4.211
Kurtosis	-.023	-.256	-.165	-.414	-.210	-.197
S E Kurt	.199	.165	.199	.165	.199	.165
Skewness	-.190	-.070	.260	.202	-.240	-.076
S E Skew	.100	.083	.100	.083	.100	.083
Range	9.000	9.000	13.000	13.000	11.000	11.000
Minimum	.000	.000	.000	.000	.000	.000
Maximum	9.000	9.000	13.000	13.000	11.000	11.000
Sum	2677.0	4156.0	3249.0	5387.0	3691.0	5536.0
Valid Cases	601	872	601	872	601	872

Table 6

Recommended Percentiles for Delphi Matrix Sub-Tests Raw Scores

	Analysis		Evaluation		Inference	
	PRE-ANAL	POST-ANAL	PRE-EVAL	POST-EVAL	PRE-INFR	POST-INFR
Number Correct						
0	1	0	1	0	1	0
1	4	2	5	3	2	1
2	10	7	11	8	3	3
3	26	21	23	16	11	8
4	49	42	37	28	21	18
5	75	69	52	42	37	34
6	92	87	70	56	55	54
7	98	97	81	69	73	71
8	99	99	89	81	88	84
9	99	99	94	88	97	94
10			98	94	99	98
11			99	97	99	98
12			99	99		
13			99	99		
14			99	99		

Deductive and Inductive Reasoning Sub-Tests

The traditional way of dividing the domain of reasoning is by distinguishing deduction and induction. These concepts, however, have become notoriously ambiguous as a result of important differences in what they denote in different disciplines. Even the notion that the one "goes from general to specific" and the other from "specific to general" has been discredited both theoretically and by counter-examples from the days of Russell and Whitehead on to the present. It is truly regrettable that this dysfunctional, nineteenth century notion can still be found in some recent methodology texts. If anything, however, this alerts us to be suspicious of any suggestion that the inductive/deductive distinction is well understood or even similarly understood across academia. Concern about this ambiguity explains why the words "deduction" and "induction" appear nowhere in the CCTST.

However, since the difference between deductive reasoning and inductive reasoning, as that distinction is understood among logicians, is both powerful and useful, the CCTST offers sub-scores in each. Logicians draw this distinction on the basis of the purported logical strength of the inference. If the assumed truth of the premises purportedly necessitates the truth of conclusion, then the argument is classified as deductive. Not only do traditional syllogisms fall within this category, but algebraic, geometric, and set-theoretical proofs in mathematics (including "mathematical induction") also represent paradigm examples of deduction. Instantiation of universalized propositions is deductive, as are inferences based on such principles as transitivity, reflexivity and identity. In the case of valid deductive arguments, it is not logically possible for the conclusion to be false and all the premises to be true.

By contrast, if an argument's conclusion is purportedly warranted, but not necessitated, by the assumed truth of its premises, logicians would consider the argument inductive. Scientific confirmation and experimental disconfirmation are examples of inductive reasoning. The day to day inferences which lead us to infer that in familiar situations things are most likely to occur or to have been caused as we have come to expect are inductions.[6] Statistical inferences are inductive, even if the inference is the prediction of an extremely probable specific (rain today) based on general principles (meteorological laws) and a given set of observations. Inference used to inform judgment by reference to perceived similarities or applications of examples, precedents, or relevant cases, such as is typical of legal reasoning, is inductive. Also inductive is that common and powerfully persuasive -- even if logically suspicious -- tool of everyday dialogue, analogical reasoning. In the case of a strong inductive argument it is unlikely or improbable that the conclusion would actually be false and all the premises true, but it is logically possible that it might.[7]

Thirty of the items on the CCTST can be readily classified as requiring the proper application of either deductive reasoning or inductive reasoning for the designated answer to be selected.[8] The CCTST, thus, offers sub-scores on deductive reasoning and inductive reasoning, as those two terms were described above. Table 7 indicates the correlations between the two sub-score and between each and the overall CCTST pretest and posttest scores. Table 8 displays statistical data for each of the two sub-tests and Table 9 indicates the recommended percentile rankings for each in both the pretest and posttest contexts using raw scores. Discretion is again recommended in the use of sub-score results and percentile rankings. Use should be restricted to diagnostic purposes, or to summative evaluations of modes of instruction, CT programs or aggregate groups of students, in contrast to summative evaluations of individuals.

Table 7

Correlations of Traditional Matrix Sub-Tests

	Deductive Reasoning		Inductive Reasoning	
	PRE-DEDU	POST-DEDU	PRE-INDU	POST-INDU
PRETEST	.8092 n= 601 p=.000	.5419 n= 412 p=.000	.7828 n= 601 p=.000	.5018 n= 412 p=.000
POSTTEST	.5602 n= 319 p=.000	.8436 n= 872 p=.000	.4218 n= 319 p=.000	.7947 n= 872 p=.000
PRE-DEDU	1.0000	.5861 n= 318 p=.000	.3219 n= 601 p=.000	.3486 n= 318 p=.000
POST-DEDU	above	1.0000	.2641 n= 318 p=.000	.3971 n= 872 p=.000
PRE-INDU	above	above	1.0000	.4589 n= 318 p=.000
POST-INDU	above	above	above	1.0000

Table 8

Statistical Analyses of Traditional Matrix Sub-Tests

	Deductive Reasoning		Inductive Reasoning	
	PRE-DEDU	POST-DEDU	PRE-INDU	POST-INDU
Mean	7.689	8.369	6.512	7.018
Std Err	.109	.096	.103	.086
Median	8.000	8.000	7.000	7.000
Mode	8.000	9.000	7.000	8.000
Std Dev	2.660	2.827	2.533	2.550
Variance	7.078	7.992	6.417	6.505
Kurtosis	-.428	-.441	-.241	-.505
S E Kurt	.199	.165	.199	.165
Skewness	.020	.167	-.016	-.093
S E Skew	.100	.083	.100	.083
Range	15.000	15.000	14.000	12.000
Minimum	.000	1.000	.000	1.000
Maximum	15.000	16.000	14.000	13.000
Sum	4621.0	7298.0	3914.0	6120.0
Valid Cases	601	872	601	872

Table 9

Recommended Percentiles for Traditional Matrix Sub-Tests Raw Scores

Number Correct	Deductive Reasoning PRE-DEDU	Deductive Reasoning POST-DEDU	Inductive Reasoning PRE-INDU	Inductive Reasoning POST-INDU
0	0	0	1	0
1	1	0	2	1
2	2	1	7	5
3	6	3	13	9
4	12	7	22	17
5	22	17	34	28
6	33	28	48	42
7	47	40	65	56
8	62	53	79	71
9	74	67	88	82
10	85	77	94	91
11	92	85	98	96
12	97	92	99	99
13	99	96	99	99
14	99	98	99	99
15	99	99		
16	99	99		

Critical Thinking Dispositions

The CCTST is designed to assess CT skills, however, the proper exercise of those skills presupposes certain crucial CT dispositions. Indeed, the CCTST includes items constructed with distractors (wrong answers) hypothesized to be more attractive to persons who do not possess the appropriate CT dispositions. Items 5, 9, 19, 20, 24-33, for example, include distractors intended to be attractive to persons who lack the dispositions identified in the Delphi study under the category of "approaches to specific issues, questions or problems," (Facione, 1990 a). Specifically, these dispositions include:

* clarity in stating the question or concern,
* orderliness in working with complexity,

* diligence in seeking relevant information,
* reasonableness in selecting and applying criteria,
* care in focusing attention on the concern at hand,
* persistence through difficulties are encountered, and
* precision to the degree permitted by the subject and the circumstances.

Likewise, items 20, 24-26, and 32-34 include other distractors which, it is hypothesized, are more likely to be selected by persons who have not developed certain of the dispositions which the Delphi research classifies under the heading of "approaches to life and living in general." Those related to the CCTST include:

* trust in the processes of reasoned inquiry,
* open-mindedness regarding divergent world views,
* flexibility in considering alternatives and opinions,
* understanding of the opinions of other people,
* fair-mindedness in appraising reasoning, and
* prudence in suspending, making, or altering judgments.

An interesting extension of this research would be to cluster such CCTST items into a sub-test on "CT-dispositions." The designated responses to items on such a sub-test would be all those choices which, it would be hypothesized, might be selected by students who approach the item with the requisite CT dispositions regardless of whether they apply their CT skills correctly. Or, in other words, the "wrong" answers to items on such a sub-test would be those distractors which would most likely be selected only by people who are hypothesized <u>not</u> to have appropriate CT dispositions. Naturally, to fully validate such a sub-test it would be necessary to conduct the kind of interviews which Steven Norris (1989) describes in his research regarding construct validation.

Conclusion

The CCTST offers three sub-tests conceptualized in terms of the Delphi study, Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction. These three sub-tests, "Analysis," "Evaluation," and "Inference," correlate strongly with each other and with the overall CCTST, used as either a pretest or a posttest. The same is true of the two CCTST sub-tests which divide along the traditional matrix, "Deductive Reasoning" and "Inductive Reasoning." Recommended percentile rankings for raw scores on each of the five and for the CCTST overall -- used either as pretests or posttests -- have been developed on the basis of analyses of 1673 test forms completed by representative samples of college students during the 1989/90 academic year at a comprehensive state university.

Partial Bibliography

Annis, David B. and Annis, Linda, "An Empirical Study of the Impact of Philosophy on Students' CT Ability, " *Teaching Philosophy* v3, pp 145-52, 1980.

Arons, Arnold B, "CT and the Baccalaureate Curriculum," *Liberal Education*, v71, n2, Summer 1985.

Blatz, Charles V., "Contextualism and CT: Programmatic Investigations," *Educational Theory* v39, n2, 1989.

Brody, Boruch A., "Glossary of Logical Terms," *Encyclopedia of Philosophy*, Paul Edwards (Ed.), Macmillan, New York, 1967, v5, p62, p66.

Ennis, Robert H., "Problems in Testing Informal Logic CT Reasoning Ability," *Informal Logic*,, v6. n1, p3-9, 1984.

_____, "A Bibliography of Testing CT," *CT News*, Center for the Reasoning Arts, CSU Sacramento, v6, n1, Sept.-Oct. 1987.

Ennis, Robert H., and Norris, Stephen P., "CT Testing and Other CT Evaluation: Status, Issues, and Needs," in *Issues in Evaluation*, Algina, James (Eds.), Ablex Press, New York, NY, 1988.

Facione, Peter A., "Assessing Inference Skills," *ERIC Clearinghouse on Tests, Measurement, and Evaluation*, Document number: TM 012917, Mar. 1989.

_____, (a) *Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction*, California Academic Press, Millbrae CA, 1990; ERIC Clearinghouse on Tests, Measurement, and Evaluation, Doc. No: TM 014423, Feb. 1990.

_____, (b) "Strategies for Multiple Choice CT Assessment," in *CT at Colleges and Universities*, David Hitchcock, (Ed.), Vale Press, Newport News, VA, 1990, forthcoming.

_____, (c) "CCTST Technical Report #1 -- **Experimental Validation and Content Validity**," California Academic Press, Millbrae, CA, Nov. 1990. (ERIC Doc. No: TM 015818).

_____. (d) "CCTST Technical Report #2 -- **Factors Predictive of CT Skills,**" California Academic Press, Millbrae, CA, Nov. 1990.

_____, (e) "CCTST Technical Report #3 -- **Gender, Ethnicity, Academic Major, CT Self-Esteem and the CCTST,**" California Academic Press, Millbrae, CA, Dec. 1990.

_____, (f) "**California Critical Thinking Skills Test: College Level,**" *California Academic Press*, 217 La Cruz, Millbrae, CA, 94030, Dec. 1990.

Kurfiss, Joanne G. *Critical Thinking: Theory, Research, Practice, and Possibilities* ASHE-ERIC Higher Education Report Number 2, Washington DC, ASHE, 1988

McPeek, John, *Critical Thinking and Education*, St. Martin's Press, New York, NY, 1981.

Norris, Stephen P. "Evaluating CT Ability," *History and Social Science Teacher*, v21, n3, p135-146, Spr. 1986.

_____, "Verbal Reports of Thinking and Multiple-Choice CT Test Design," *Technical Report No. 447*, Champaign, IL: Center for the Study of Reading, University of Illinois, (ERIC Doc. No: ED302826.)

_____, "Effect of Eliciting Verbal Reports of Thinking on CT Test Performance," *Journal of Educational Measurement*, v27, n1, 1990.

Norris, Stephen P., and Ennis, Robert H., *Evaluating CT*, Midwest Publications, Pacific Grove, CA, 1989.

Pecorino, Philip, "CT Bibliography," *Newsletter on Teaching Philosphy*, American Philsophical Association, p18, Summer 1987.

Resnick, L. W., *Education and Learning to Think*, National Academy Press, 1987.

Ruggiero, V. R., *Teaching Thinking Across the Curriculum*, Harper and Row, New York, NY, 1988.

Siegel, Harvey, *Educating Reason: Rationality, CT, and Education*, Routledge, London, 1988.

Sternberg, Robert J., "CT: Its Nature, Measurement, and Improvement," National Institute of Education, Washington, DC, 1986.

Endnotes

[1] Control group students group means were not statistically significantly different than the control group pretest mean, hence their scores were used to supplement the size of the sample used to establish these pretest norms.

[2] Relevant data on SAT scores and college GPA for the norm groups was presented in TR #2, "Factors Predictive of CT Skills".

[3] Using the names of the core CT skills identified in the Delphi research, items 1-5 target interpretation, items 6-9 analysis, items 10-13 evaluation, items 14-24 inference, and items 25-34 explanation.

[4] Data on possible differences by gender, ethnicity, or academic major on the various sub-tests is yet to be analyzed. Likewise the relationships of various sub-test scores to other indicators such as SAT or GPA has yet to be determined.

[5] Statistical significance reported here ($p < .001$) was obtained using the one-tailed test for the significance of Pierson-rho.

[6] Sherlock Holmes used induction but called it deduction.

[7] The Encyclopedia of Philosophy, which is authoritative in such matters, defines "deduction" as "a form of inference such that in a valid deductive argument the joint assertion of the presmises and the denial of the conclusion is a contradiction." The reference goes on to contrast inferences of the deductive variety with those in which that joint assertion would not be contradictory and includes inductive inferences as among this latter group.

[8] The sixteen deductive items are: 1, 2, 4, 5, 11, 12, 13, 14, 15, 16, 17, 18, 19, 22, 23, and 30. The fourteen inductive items are: 9, 10, 20, 21, 24, 25, 26, 27, 28, 29, 31, 32, 33, and 34. The excluded items are: 3, 6, 7, and 8.